

#9



OIPE

RAW SEQUENCE LISTING

DATE: 05/07/2002

PATENT APPLICATION: US/09/866,570A

TIME: 11:40:38

Input Set : N:\Crif3\RULE60\09866570A.raw

Output Set: N:\CRF3\05072002\I866570A.raw

1 <110> APPLICANT: Croteau, Rodney et al.
2 <120> TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway
3 <130> FILE REFERENCE: 53679
4 <140> CURRENT APPLICATION NUMBER: 09/866,570A
5 <141> CURRENT FILING DATE: 2001-05-25
7 <150> PRIOR APPLICATION NUMBER: US/09/457,046B
8 <151> PRIOR FILING DATE: 1999-12-07

10 <160> NUMBER OF SEQ ID NOS: 74
11 <170> SOFTWARE: PatentIn Ver. 2.1

13 <210> SEQ ID NO: 1

14 <211> LENGTH: 920

15 <212> TYPE: DNA

16 <213> ORGANISM: Taxus cuspidata

17 <400> SEQUENCE: 1

18 atgttggtct attatcccc ttttgctggg cgcctcagag agacagaaaa tggggatctg 60
19 gaagtggat gcacagggga ggggtgctatg tttttggaag ccatggcaga caatgagctg 120
20 tctgtgttg gagattttga tgacagcaat ccatcatttc agcagctact tttttcgctt 180
21 ccactcgata ccaatttcaa agacctctct cttctggttg ttcaggtaac tcgttttaca 240
22 tgtggaggct ttgttggttg agtgagtttc caccatggtg tatgtgatgg tcgaggagcg 300
23 gcccaatttc ttaaagggtt ggcagaaatg gcacggggag aggttaagct ctcatgggaa 360
24 ccaatatgga atatggaact agtgaagctt gatgacccta aatacctcca attttttcac 420
25 tttgaattcc tacgagcgcc ttcaattggt gagaaaattg ttcaaacata ttttattata 480
26 gatttgagga ccataaatta tatcaaaca tctgttatgg aagaatgtaa agaattttgc 540
27 tcttcattcg aagttgcatc agcaatgact tggatagcaa ggacaagagc ttttcaaatt 600
28 ccagaaagtg agtacgtgaa aattctcttc ggaatggaca tgaggaactc atttaatccc 660
29 cctcttccaa gcggatacta tggtaactcc attggtaccg catgtgcagt ggataatggt 720
30 caagacctct taagtggatc tcttttgctg gctataatga ttataaagaa atcaaaggtc 780
31 tcttttaaatg ataatttcaa gtcaagagct gtggtgaagc catctgaatt ggatgtgaat 840
32 atgaatcatg aaaacgtagt tgcatttgct gattggagcc gattgggatt tgatgaagtg 900
33 gattttggct gggggaaacc 920

35 <210> SEQ ID NO: 2

36 <211> LENGTH: 306

37 <212> TYPE: PRT

38 <213> ORGANISM: Taxus cuspidata

39 <400> SEQUENCE: 2

40 Met Leu Val Tyr Tyr Pro Pro Phe Ala Gly Arg Leu Arg Glu Thr Glu
41 1 5 10 15
42 Asn Gly Asp Leu Glu Val Glu Cys Thr Gly Glu Gly Ala Met Phe Leu
43 20 25 30
44 Glu Ala Met Ala Asp Asn Glu Leu Ser Val Leu Gly Asp Phe Asp Asp
45 35 40 45
46 Ser Asn Pro Ser Phe Gln Gln Leu Leu Phe Ser Leu Pro Leu Asp Thr
47 50 55 60

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/866,570A

DATE: 05/07/2002

TIME: 11:40:38

Input Set : N:\Crif3\RULE60\09866570A.raw

Output Set: N:\CRF3\05072002\I866570A.raw

```

48   Asn Phe Lys Asp Leu Ser Leu Leu Val Val Gln Val Thr Arg Phe Thr
49       65                               70                               75                               80
50   Cys Gly Gly Phe Val Val Gly Val Ser Phe His His Gly Val Cys Asp
51                               85                               90                               95
52   Gly Arg Gly Ala Ala Gln Phe Leu Lys Gly Leu Ala Glu Met Ala Arg
53                               100                              105                              110
54   Gly Glu Val Lys Leu Ser Leu Glu Pro Ile Trp Asn Met Glu Leu Val
55                               115                              120                              125
56   Lys Leu Asp Asp Pro Lys Tyr Leu Gln Phe Phe His Phe Glu Phe Leu
57       130                               135                               140
58   Arg Ala Pro Ser Ile Val Glu Lys Ile Val Gln Thr Tyr Phe Ile Ile
59       145                               150                               155                               160
60   Asp Leu Glu Thr Ile Asn Tyr Ile Lys Gln Ser Val Met Glu Glu Cys
61                               165                              170                              175
62   Lys Glu Phe Cys Ser Ser Phe Glu Val Ala Ser Ala Met Thr Trp Ile
63                               180                              185                              190
64   Ala Arg Thr Arg Ala Phe Gln Ile Pro Glu Ser Glu Tyr Val Lys Ile
65       195                              200                              205
66   Leu Phe Gly Met Asp Met Arg Asn Ser Phe Asn Pro Pro Leu Pro Ser
67       210                              215                              220
68   Gly Tyr Tyr Gly Asn Ser Ile Gly Thr Ala Cys Ala Val Asp Asn Val
69       225                              230                              235                              240
70   Gln Asp Leu Leu Ser Gly Ser Leu Leu Arg Ala Ile Met Ile Ile Lys
71                               245                              250                              255
72   Lys Ser Lys Val Ser Leu Asn Asp Asn Phe Lys Ser Arg Ala Val Val
73                               260                              265                              270
74   Lys Pro Ser Glu Leu Asp Val Asn Met Asn His Glu Asn Val Val Ala
75       275                              280                              285
76   Phe Ala Asp Trp Ser Arg Leu Gly Phe Asp Glu Val Asp Phe Gly Trp
77       290                              295                              300
78   Gly Lys
79       305
81 <210> SEQ ID NO: 3
82 <211> LENGTH: 920
83 <212> TYPE: DNA
84 <213> ORGANISM: Taxus cuspidata
85 <400> SEQUENCE: 3
86   atgctggtct attatccccc ttttgctgga aggctgagaa acacagaaaa tggggaactt 60
87   gaagtggagt gcacagggga ggggtgccgtc tttgtggaag ccatggcgga caacgacctt 120
88   tcagtattac aagatttcaa tgagtacgat ccatcatttc agcagctagt tttttatctt 180
89   ccagaggatg tcaatattga ggacctccat cttctaactg ttcaggtaac tcgttttaca 240
90   tgtgggggat ttgttgtggg cacaagattc caccatagtg tgtctgatgg aaaaggaatc 300
91   ggccagttac ttaaaggcat gggagaaatg gcaagggggg agtttaagcc ctccttagaa 360
92   ccaatatgga atagagaaat ggtgaagcct gaagacatta tgtacctcca gtttgatcac 420
93   tttgatttca tacaccacc tcttaatctt gagaagtcta ttcaagcatc tatggtaata 480
94   agcttggaga gaataaatta tatcaaacga tgcattgatg aagaatgcaa agaatttttt 540
95   tctgcatttg aagttgtagt agcattgatt tggctagcaa ggacaaagtc ttttcgaatt 600
96   ccaccaatg agtatgtgaa aattatcttt ccaatcgaca tgaggaattc atttgactcc 660
97   cctcttccaa agggatacta tggtaatgct attggtaatg catgtgcaat ggataatgct 720

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/866,570A

DATE: 05/07/2002

TIME: 11:40:38

Input Set : N:\Crf3\RULE60\09866570A.raw

Output Set: N:\CRF3\05072002\I866570A.raw

```

98      aaagacctct taaatggatc tcttttatat gctctaagtc ttataaagaa atcaaagttt 780
99      gcttttaaagc agaatttcaa atcaagaatc ttgacaaaac catctgcatt agatgcgaat 840
100     atgaagcatg aaaatgtagt cggatgtggc gattggagga atttgggatt ttatgaagca 900
101     gatttcggct ggggcaaacc                                     920
103 <210> SEQ ID NO: 4
104 <211> LENGTH: 306
105 <212> TYPE: PRT
106 <213> ORGANISM: Taxus cuspidata
107 <400> SEQUENCE: 4
108     Met Leu Val Tyr Tyr Pro Pro Phe Ala Gly Arg Leu Arg Asn Thr Glu
109           1             5             10             15
110     Asn Gly Glu Leu Glu Val Glu Cys Thr Gly Glu Gly Ala Val Phe Val
111           20             25             30
112     Glu Ala Met Ala Asp Asn Asp Leu Ser Val Leu Gln Asp Phe Asn Glu
113           35             40             45
114     Tyr Asp Pro Ser Phe Gln Gln Leu Val Phe Tyr Leu Pro Glu Asp Val
115           50             55             60
116     Asn Ile Glu Asp Leu His Leu Leu Thr Val Gln Val Thr Arg Phe Thr
117           65             70             75             80
118     Cys Gly Gly Phe Val Val Gly Thr Arg Phe His His Ser Val Ser Asp
119           85             90             95
120     Gly Lys Gly Ile Gly Gln Leu Leu Lys Gly Met Gly Glu Met Ala Arg
121           100            105            110
122     Gly Glu Phe Lys Pro Ser Leu Glu Pro Ile Trp Asn Arg Glu Met Val
123           115            120            125
124     Lys Pro Glu Asp Ile Met Tyr Leu Gln Phe Asp His Phe Asp Phe Ile
125           130            135            140
126     His Pro Pro Leu Asn Leu Glu Lys Ser Ile Gln Ala Ser Met Val Ile
127           145            150            155            160
128     Ser Leu Glu Arg Ile Asn Tyr Ile Lys Arg Cys Met Met Glu Glu Cys
129           165            170            175
130     Lys Glu Phe Phe Ser Ala Phe Glu Val Val Val Ala Leu Ile Trp Leu
131           180            185            190
132     Ala Arg Thr Lys Ser Phe Arg Ile Pro Pro Asn Glu Tyr Val Lys Ile
133           195            200            205
134     Ile Phe Pro Ile Asp Met Arg Asn Ser Phe Asp Ser Pro Leu Pro Lys
135           210            215            220
136     Gly Tyr Tyr Gly Asn Ala Ile Gly Asn Ala Cys Ala Met Asp Asn Val
137           225            230            235            240
138     Lys Asp Leu Leu Asn Gly Ser Leu Leu Tyr Ala Leu Met Leu Ile Lys
139           245            250            255
140     Lys Ser Lys Phe Ala Leu Asn Glu Asn Phe Lys Ser Arg Ile Leu Thr
141           260            265            270
142     Lys Pro Ser Ala Leu Asp Ala Asn Met Lys His Glu Asn Val Val Gly
143           275            280            285
144     Cys Gly Asp Trp Arg Asn Leu Gly Phe Tyr Glu Ala Asp Phe Gly Trp
145           290            295            300
146     Gly Lys
147           305

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/866,570A

DATE: 05/07/2002

TIME: 11:40:38

Input Set : N:\Crif3\RULE60\09866570A.raw

Output Set: N:\CRF3\05072002\I866570A.raw

```

149 <210> SEQ ID NO: 5
150 <211> LENGTH: 903
151 <212> TYPE: DNA
152 <213> ORGANISM: Taxus cuspidata
153 <400> SEQUENCE: 5
154      ttttatccgt ttgcggggcg gctcagaaat aaagaaaatg ggggaacttga agtggagtgc 60
155      acagggcagg gtgttctgtt tctggaagcc atggccgaca gcgaccttc agtcttaaca 120
156      gatctggatg actacaagcc atcgttttcag cagttgattt tttctctacc acaggataca 180
157      gatattgagg atctccatct cttgattgtt caggtaactc gttttacatg tgggggtttt 240
158      gttgtgggag cgaatgtgta tagtagtgta tgtgatgcaa aaggatttgg ccaatttctt 300
159      caaggtatgg cagagatggc gagaggagag gttaagccct cgattgaacc gatatggaat 360
160      agagaactgg tgaagccaga acattgtatg cccttccgga tgagtcatct tcaaattata 420
161      cacgcacctc tgatcgagga gaaatttgtt caaacatctc ttgttataaa ctttgagata 480
162      ataaatcata tcagacaacg gatcatggaa gaatgtaaag aaagtttctc ttcatttgaa 540
163      attgtagcag cattggtttg gctagcaaag ataaaggctt ttcaaattcc acatagttag 600
164      aatgtgaagc ttctttttgc aatggactta aggagatcat ttaatccccc tcttccacat 660
165      ggatactatg gcaatgcctt cggtattgca tgtgcaatgg ataatgtcca tgacctttta 720
166      agtggatctc ttttgcgcgc tataatgata ataaagaaat caaagttctc tttacacaaa 780
167      gaactcaact caaaaaccgt gatgagcccg tctgtagtag atgtcaatac gaagttcgaa 840
168      gatgtagttt caattagtga ctggaggcag tctatatatt atgaagtgga ctttggttgg 900
169      ggc
170
171 <210> SEQ ID NO: 6
172 <211> LENGTH: 301
173 <212> TYPE: PRT
174 <213> ORGANISM: Taxus cuspidata
175 <400> SEQUENCE: 6
176      Phe Tyr Pro Phe Ala Gly Arg Leu Arg Asn Lys Glu Asn Gly Glu Leu
177      1          5          10          15
178      Glu Val Glu Cys Thr Gly Gln Gly Val Leu Phe Leu Glu Ala Met Ala
179      20          25          30
180      Asp Ser Asp Leu Ser Val Leu Thr Asp Leu Asp Asp Tyr Lys Pro Ser
181      35          40          45
182      Phe Gln Gln Leu Ile Phe Ser Leu Pro Gln Asp Thr Asp Ile Glu Asp
183      50          55          60
184      Leu His Leu Leu Ile Val Gln Val Thr Arg Phe Thr Cys Gly Gly Phe
185      65          70          75          80
186      Val Val Gly Ala Asn Val Tyr Ser Ser Val Cys Asp Ala Lys Gly Phe
187      85          90          95
188      Gly Gln Phe Leu Gln Gly Met Ala Glu Met Ala Arg Gly Glu Val Lys
189      100         105         110
190      Pro Ser Ile Glu Pro Ile Trp Asn Arg Glu Leu Val Lys Pro Glu His
191      115         120         125
192      Cys Met Pro Phe Arg Met Ser His Leu Gln Ile Ile His Ala Pro Leu
193      130         135         140
194      Ile Glu Glu Lys Phe Val Gln Thr Ser Leu Val Ile Asn Phe Glu Ile
195      145         150         155         160
196      Ile Asn His Ile Arg Gln Arg Ile Met Glu Glu Cys Lys Glu Ser Phe
197      165         170         175
198      Ser Ser Phe Glu Ile Val Ala Ala Leu Val Trp Leu Ala Lys Ile Lys

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/866,570A

DATE: 05/07/2002

TIME: 11:40:38

Input Set : N:\Crif3\RULE60\09866570A.raw

Output Set: N:\CRF3\05072002\I866570A.raw

```

199          180          185          190
200      Ala Phe Gln Ile Pro His Ser Glu Asn Val Lys Leu Leu Phe Ala Met
201          195          200          205
202      Asp Leu Arg Arg Ser Phe Asn Pro Pro Leu Pro His Gly Tyr Tyr Gly
203          210          215          220
204      Asn Ala Phe Gly Ile Ala Cys Ala Met Asp Asn Val His Asp Leu Leu
205          225          230          235          240
206      Ser Gly Ser Leu Leu Arg Ala Ile Met Ile Ile Lys Lys Ser Lys Phe
207          245          250          255
208      Ser Leu His Lys Glu Leu Asn Ser Lys Thr Val Met Ser Pro Ser Val
209          260          265          270
210      Val Asp Val Asn Thr Lys Phe Glu Asp Val Val Ser Ile Ser Asp Trp
211          275          280          285
212      Arg Gln Ser Ile Tyr Tyr Glu Val Asp Phe Gly Trp Gly
213          290          295          300
215 <210> SEQ ID NO: 7
216 <211> LENGTH: 908
217 <212> TYPE: DNA
218 <213> ORGANISM: Taxus cuspidata
219 <400> SEQUENCE: 7
220      ttctaccctg ttgcagggcg gctcagaaat aaagaaaatg gggaacttga agtggagtgc 60
221      acagggcagg gtgttctgtt tctggaagcc atggctgaca gcgacgtttc agtcttaaca 120
222      gatctggaag actacaatcc atcgtttcag cagttgcttt tttctctacc acaggataca 180
223      gatattgagg acctccatct cttgattgtt caggtgactc actttacatg tggggatttt 240
224      gttgtgggag cgaatgttta tggtagtgta tgtgacggaa aaggatttgg ccagtttctt 300
225      caaggtatgg cggagatggc gagaggagag gttaagccct cgattgaacc gatatggaat 360
226      agagaactgg tgaagccaga agatttaatg gccctccacg tggatcatct tcgaattata 420
227      cacacacctc taatcgagga gaaatttggt caaacatctc ttgttataaa ctttgagata 480
228      ataaatcata tcagacgatg catcatggaa gaatgtaaag aaagtttctc ttcattcgaa 540
229      attgtagcag cattggtttg gctagcaaag ataaaagctt ttcgaattcc acatagttag 600
230      aatgtgaaga ttctctttgc aatggacgtg aggagatcat ttaagcccc tcttccaaag 660
231      ggatactatg gcaatgccta tggatttgca tgtgcaatgg ataatgtcca ggatcttcta 720
232      agtggatctc ttttgcatgc tataatgatc ataaagaaat caaagttctc tttacacaaa 780
233      aaaatcaact caaaaactgt gatgagcccc tctccattag acgtcaatat gaagtttgaa 840
234      aatgtagttt caattactga ttggaggcat tctaaatatt atgaagtaga cttcgggtgg 900
235      ggtaaacc
237 <210> SEQ ID NO: 8
238 <211> LENGTH: 302
239 <212> TYPE: PRT
240 <213> ORGANISM: Taxus cuspidata
241 <400> SEQUENCE: 8
242      Phe Tyr Pro Phe Ala Gly Arg Leu Arg Asn Lys Glu Asn Gly Glu Leu
243          1          5          10          15
244      Glu Val Glu Cys Thr Gly Gln Gly Val Leu Phe Leu Glu Ala Met Ala
245          20          25          30
246      Asp Ser Asp Val Ser Val Leu Thr Asp Leu Glu Asp Tyr Asn Pro Ser
247          35          40          45
248      Phe Gln Gln Leu Leu Phe Ser Leu Pro Gln Asp Thr Asp Ile Glu Asp
249          50          55          60

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/866,570A

DATE: 05/07/2002
TIME: 11:40:39

Input Set : N:\Crf3\RULE60\09866570A.raw
Output Set: N:\CRF3\05072002\I866570A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:18; Xaa Pos. 164
Seq#:34; N Pos. 3,6,9,18
Seq#:35; N Pos. 9,12,18
Seq#:36; N Pos. 9,15,18
Seq#:37; N Pos. 9,12,15,18
Seq#:38; N Pos. 9,15
Seq#:48; Xaa Pos. 2,3,4

VERIFICATION SUMMARY

DATE: 05/07/2002

PATENT APPLICATION: US/09/866,570A

TIME: 11:40:39

Input Set : N:\Crf3\RULE60\09866570A.raw

Output Set: N:\CRF3\05072002\I866570A.raw

L:606 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18 after pos.:160
L:1066 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1069 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:34
L:1070 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:34 after pos.:0
L:1078 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1081 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:35
L:1082 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:35 after pos.:0
L:1090 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1093 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:36
L:1094 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:36 after pos.:0
L:1102 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1105 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:37
L:1106 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:37 after pos.:0
L:1114 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1117 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:38
L:1118 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:38 after pos.:0
L:1286 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1289 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:48
L:1290 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:48 after pos.:0